Working Papers in Economics & Finance
2020-14

Impact evaluation, social performance assessment and standardisation: reflections from microfinance evaluations in Pakistan and Zimbabwe

Joana Silva Afonso, University of Portsmouth

Portsmouth Business School
https://www.port.ac.uk/about-us/structure-and-governance/organisational-structure/our-academic-structure/faculty-of-business-and-law/portsmouth-business-school

Impact evaluation, social performance assessment and standardisation: reflections from microfinance evaluations in Pakistan and Zimbabwe[1]

Joana Silva Afonso, University of Portsmouth

Abstract

Heterogeneity associated with the microfinance and financial inclusion sector discourages the application of 'one-fits-all' evaluation models. However, under certain conditions, there can be advantages in adopting a 'common project approach' to evaluation. This paper is based on my participation as academic consultant on a project led by the crowdfunding platform Lendwithcare, which aimed to assess the outcomes at client level of two microcredit programmes developed by its partner microfinance institutions in Pakistan and Zimbabwe. Applying qualitative methodologies, including participant-observation and interviews, I analyse the conditions in which some degree of standardisation is feasible, its advantages and limitations and how this experience can add to the knowledge on impact evaluation and social performance assessment in the sector. In doing so, I establish a parallel with previous evaluation projects in the sector, namely the AIMS, Imp-Act and Microfinance for Decent Work projects.

The findings show that the approach followed in the Lendwithcare project contributed to change mind-sets regarding evaluation and trigger the process of social performance management in the participating institutions, which had incipient experience in measuring social performance and were not familiar with the evaluation process. It attracted also the attention of other Lendwithcare partner MFIs whose managers showed interest in replicating the process in their institutions.

1. Introduction

In this new age of digital finance and the wonders and dangers associated with its spreading in the developing countries, does it still matter to look into the smaller microfinance institutions (MFIs) with a strong social focus for which traditional microfinance tools, particularly microcredit for businesses, remains the main product offered to poor and marginalised populations? Do they still have a role in promoting financial inclusion in the regions where they operate? If one takes as indicative the result of the discussion in the closing session of the European Microfinance Week 2018 on the topic "Is there room left for the 'little guy'?" practitioners seem to think so.[2]

What about research? Has 'everything' be written on the topic and researchers should move on to new topics? Considering that this is not the case, this paper analyses the implementation of

[1] Acknowledgments: I would like to extend my gratitude to our research partners, Lendwithcare, which promoted the impact evaluation, and Akhuwat Islamic Microfinance in Pakistan and Thrive Microfinance in Zimbabwe, particularly t the local evaluation teams involved in the project.

[2] The session can be viewed at: https://registration.european-microfinance-week.eu/ehome/emw2018/streaming/

an evaluation project by a social investor in two MFIs in different contexts, which have in common the focus on microcredit as main tool to achieve social goals and, for different reasons, incipient efforts to measure their social performance.

The evaluation project tried to conciliate two purposes: an 'overall summative judgment', aiming to obtain 'credible' evidence of the social outcomes of the programmes supported (at the eyes of their funders) and a 'formative learning and improvement' process, building capacity within the MFIs to implement social performance measurement and management systems.[3] In order to increase the credibility of the results and guarantee a better methodological approach, given the resources available and the learning perspective adopted, Lendwithcare (LWC),[4] the leader of the project looked for an academic consultant to be partner in the project.

It was the participation of the University of Portsmouth as academic consultant in the LWC evaluation project that allowed for the development of this research into the advantages and limitations of the 'common project approach' followed in this project. One of the aims of the paper is, thus, to keep alive the discussion regarding how different perspectives on the evaluation of outcomes and impact can contribute to increase accountability and transparency in the sector and improve the microfinance and financial inclusion programmes implemented. This being true particularly for Tier 2 and 3 MFIs,[5] which are more unlikely to be the object of impact evaluations fulfilling the highest academic standards of rigour.

The research follows the footsteps of previous research associated with the Imp-Act and Microfinance for Decent Work projects (ILO, 2015; Copestake & Simanowitz, 2005), and further behind, the Assessing the Impact of Microenterprise Services (AIMS) project (Barnes & Sebstad, 2000), exploring similarities and differences between the approaches and what the accumulated knowledge from these projects brings into the sector. It is, however, important to stress that in any way I look to undermine the methodological limitations associated with the evaluations implemented in the project showcased in the paper and the previous projects. The message of the paper is that these different perspectives on (impact) evaluation can co-exist and contribute to improve the programmes, as suggested by White (2009).

The paper is organised as follows. Section 2 contextualises the research and discusses the methodological issues associated with the different perspectives on impact evaluation. Section 3 details the research design and briefly presents the LWC evaluation project. Section 4 discusses the 'common project approach' adopted, focusing on its elements of standardisation. Finally, section 5 concludes, summarising the findings and lessons learned from the project.

[3] This classification of the evaluation purposes was put forward by Patton (2008, p.139) in the framework of its utilisation-focused evaluation model. In addition to these, Patton considers 'accountability' (assessment of the effectiveness of the programme), 'monitoring', 'developmental' and 'knowledge generating' purposes.

[4] Lendwithcare is a crowdfunding platform, which supports poor entrepreneurs in the developing countries by funding business loans provided by local MFIs. It is an initiative of development NGO CARE International UK. Further information on the institution is provided in section 3 of the paper.

[5] The definition of Tier 2 and 3 MFIs used in the paper follows the proposal of the European Microfinance Platform Action Group of Investors in Tier 2/3 MFIs (e-MFP, 2013).

2. Microfinance evaluation and impact: the perspective of Tier 2 and 3 MFIs

The majority of the MFIs classified as Tier 2 and 3 are, as described by Hughes and Hutchings (2011, p.iii) with respect to development NGOs, "not set up to rigorously evaluate" their work, and particularly assess social performance. By the end of the 1990s, when microcredit experiences started gaining more visibility, Morduch (1999) pointed out that there was no incentive to evaluate the programmes. Donors/funders were more concerned with the (solid) financial performance of the MFIs, being satisfied with anecdotic cases reporting the success at client level of the microcredit programme.

This was generally the case until the mid-2000s, with the positive results reported by impact studies such as Barnes (2001), Littlefield et al. (2003) and Pitt and Khandker (1998), among others, contributing to the belief that microcredit was (among) the best development tools to alleviate poverty and empower women in the developing countries. Not that there were not dissonant voices, including Fernando (1997), Kabeer (2001), Morduch (1998), Vogelgesang (2003), and Von Pischke and Adams (1992). These authors were more moderated in regard to the outcomes of the microcredit programmes and have called attention to situations in which these outcomes could be potentially negative for the clients. These arguments were, however, muffled by mainstream thought that microcredit could be the desired magic bullet to solve some of the most challenging social problems.

The delinquency crises in different countries, including Bosnia-Herzegovina, India, Morocco, Nicaragua and Pakistan, during the 2000s (Chen et al., 2010; Guérin et al., 2015), which were intensified by the end of the decade by the world financial crisis, proved that it was not the case. This led to a change of attitude of the main actors in the sector towards the evaluation of the programmes and accountability of the institutions. Financial performance was no longer the only concern, with social performance becoming an equal priority for many donors and social investors.

In 2005, the Social Performance Task Force (SPTF) was created (SPTF, n.d.). SPTF is a network of institutions and individuals working in the sector, including donors, investors, consultants and MFIs, which has been developing advocacy work on the importance of social performance and has been working with other partners (including CERISE), to create practical mechanisms to help the MFIs assess their own performance. Among these mechanisms, the Universal Standards for Social Performance Management (USSPM) became reference guidelines for MFIs looking to assess social performance and ensure they are achieving their social mission.

At academic level, a significant number of studies described the delinquency crises and aimed to dissect the causes and responses associated with the different country scenarios (Bastiaensen et al., 2013; Bateman, 2010; Burki, 2009; Chen et al., 2010; Mader, 2013; Morvant-Roux & Roesch, 2013; Nair, 2010; Rhyne, 2001; Rozas, 2011). Linked to this, a strand of literature has focused on the causes, consequences and assessment of client over-indebtedness (Afonso et al., 2017; Bylander et al., 2018; Schicks, 2013; Schicks & Rosenberg, 2011; Gonzaléz, 2008; Guérin et al., 2013; Morvant-Roux et al., 2015; Smits, 2017), following the pioneer work of Vogelgesang (2003).

Between 2012 and 2015, the European Investment Bank funded a large research project on the crises - the 'Microfinance in Crisis' project. This project analysed different elements associated with the demand, supply and environment of the microfinance sector in five different contexts, which were affected (India, Morocco, Nicaragua) or were judged in risk of being affected (Senegal, the Dominican Republic) by delinquency crises (Guérin et al., 2015).[6] Beyond a number of research policy briefs, academic papers and a book, the project produced a crises prevention dashboard.

At the same time, the methodologies used to assess outcomes and impact have evolved, being commonly accepted that the quasi-experimental methodologies applied in the studies mentioned above most likely led to an overestimation of the positive outcomes regarding poverty reduction, as these methodologies are not sufficient to eliminate the selection bias and fungibility problems associated with the programmes. Therefore, the application of different methodologies, including RCTs - randomised control trials (Banerjee & Duflo, 2011), and mixed methods combining quantitative and qualitative methodologies (Copestake, 2012; Odell, 2015; Reichardt, 2011), has been advocated in order to obtain more robust results and make better decisions regarding the merit of the programmes.

This broadening of the methodological choices reflects different perspectives on evaluation and impact assessment. White (2009) calls attention for the use of the term 'impact' to refer to two different concepts, considering that neither of them is right or wrong, and both can be potentially useful for policy makers.

On one hand, an 'impact assessment' from an academic perspective implies the attribution (in some degree) of the changes observed in the selected outcomes to the participation in the programme. The focus is on cause-and-effect questions, and requires dealing with endogeneity and fungibility problems (Adams & Vogels, 2013; Duvendack et al., 2011; Gertler et al., 2016). This is an approach designated as 'formal impact assessment' by Copestake (2012).[7] The author suggests that in the microfinance sector this approach is frequently associated with studies conducted by external evaluators who purposely deal with the problem of attribution, often through statistical inference. As mentioned above, RCTs have been considered by many authors (Banerjee et al.; 2015; Boruch, 1997; Duflo, 2004; Van Rooyen et al., 2012), particularly positivist researchers, as the standard to be followed in impact assessment. A lively debate has been taking place within the sector (and other development areas) about the merits and limitations of this methodology (see e.g Bédécarrats, Guérin & Roubaud, 2015; Karlan, Goldberg & Copestake, 2009; Ravallion, 2012; Scriven, 2008). Copestake (2012) offers a more pluralistic view, considering like other researchers (Balkenhol, 2012; Marr & Awaworyi, 2012; Reichardt, 2011: White, 2009) that an exclusive focus on RCTs is a too narrow perspective, and that the choice of methodologies

[6] The research project was coordinated by the Université Paris I Sorbonne/IRD and involved researchers from the Centre for European Research in Microfinance (CERMi, Belgium), the University of Fribourg (Switzerland) and the Laboratoire de Statistique Appliquée à l'Analyse et la Recherche en Economie (Morocco).

[7] Copestake (2012) distinguishes between formal impact assessments, informal impact assessments and broader social science research.

should depend on the objectives and conditions of the impact evaluations. Mixed methods may be alternatively the most adequate approach, combining quantitative and qualitative methodologies, while in some cases it can also be justifiable the application of only qualitative methodologies, such as the QUIP (Wright and Copestake, 2004).

Although being important to acknowledge this academic debate on impact methodologies, this is not the focus of the paper. Rather more significant is to consider the second concept referred to by White (2009), which corresponds to a perspective generally adopted by practitioners.[8] Impact is associated with the long-term outcomes of the microfinance programmes, including wider impacts, such as changes at the community level (Rogers, 2014). In this perspective, the focus of the evaluations is on the characterisation of the changes observed in the businesses and lives of the microfinance clients and their households, with the attribution of these changes specifically to the programme evaluated not being the main priority. In the paper, my standpoint is of acceptance of the co-existence of these different perspectives, as suggested by White (2009), provided that it is clear which concept is being applied and there is transparency in the communication of its advantages and limitations. The paper findings suggest that a more nuanced perspective on evaluation and impact can encourage the Tier 2/3 MFIs to embrace social performance and evaluation as a part of their internal strategies, in line with the call of Hulme (2000) for "strengthening the impact monitoring capacity of the MFI itself" (p.93).

At the MFIs level, the need to legitimise the work of the institutions has been increasingly felt. However, for the Tier 2/3 MFIs with limited financial resources, which often do not integrate staff with specific skills in social performance and evaluation, donors and investors, such as Lendwithcare, can play an important role. They can promote transparency in the report of outcomes of the programmes and support research on impact (Simanowitz, 2003), including projects focusing on building capacity within the institutions.

In this context, it is important to highlight a few collaborative projects between academics and practitioners, which represent landmarks in the dissemination of an evaluation culture within the sector. These were the cases of the above mentioned AIMS, Imp-Act and Microfinance for Decent Work projects (Barnes and Sebstad, 2000; Copestake et al., 2005; ILO, 2015), with the lessons and results of the projects being used beyond the participating institutions; contributing for example to shape the Universal Standards of Social Performance Management (USSPM) defined by the SPTF.

These action-research projects are important references for their scale, involving multiple MFIs in different countries, and their high-profile sponsors (USAID, Ford Foundation and ILO), which contributed to the visibility and credibility of their findings. In these projects, although in different degrees, the local institutions were involved in the design and implementation of the evaluations conducted. In the case of the Imp-Act project "each (MFI) was responsible for its own work" (p.8), an approach expected to build internal capacity on evaluation and social performance and

[8] Practitioner is considered here in a wide sense, including not only the microfinance institutions and other financial service providers offering formal finance to poor and vulnerable populations, but also other actors working in the field, including donors, social investors, rating agencies, consultants, etc.

increase the commitment of the MFI managers towards accountability and transparency (Copestake and Simanowitz, 2005).

Imp-Act focused on social performance assessment mechanisms, which could be facilitators of the MFIs' decision-making process and were capable of satisfying the needs of their multiple stakeholders, having as ultimate goal the improvement of the quality of the financial services provided and their impact on poverty (Copestake and Simanowitz, 2005). The project involved 30 partner MFIs, with a central team of researchers in the UK providing "limited technical advice when requested and where appropriate" (p.6), and monitoring progress of the individual action-research programmes, which were implemented between 2001 and 2004.

In the contextualisation of their own project, Copestake and Simanowitz (2005) highlight the role of the AIMS project in the late 1990s in changing perceptions on impact evaluation in the sector by calling attention to "ways of institutionalising assessment of outreach and impact on clients within routine operations of MFIs" (p.3). This recognition is consistent with Simanowitz (2001, p.11) who describes a trend in microfinance impact assessment to "move away from donor-led impact events, towards more practitioner-focused processes".

The AIMS project, developed by the United States Aid Agency (USAID) between 1995 and 2002, had two main objectives: understanding the processes by which microenterprise services contribute to strengthen businesses and improve the well-being of the entrepreneurs and their households; and improve the capacity of USAID and its local partners to evaluate the impact of the programmes implemented (Dunn, 1999). The relevance of the project in the context of this paper relates to the three impact evaluations conducted in India, Peru and Zimbabwe under the AIMS Project Core Impact Assessment. These three evaluations were the basis for the development of specific manuals and tools (Barnes and Sebstad, 2000), which were used or served as inspiration in many of the evaluations conducted in both Imp-Act and Microfinance for Decent Work projects (Greeley, 2005; ILO, 2015).

The Microfinance for Decent Work project, an initiative of the ILO had as main goals the implementation by microfinance institutions of innovative financial and non-financial services, and its respective evaluation in terms of social impact. An important element in the project design was the expectation that experimental methodologies (RCTs) were implemented in the impact assessment stage in all settings, considering that these methodologies would increase the credibility of the findings (ILO, 2015).

The project involved initially 25 MFIs, from which 18 introduce innovations in their microfinance offer and 13 completed longitudinal evaluations.[9] The initial intention of conducting RCTS in all MFIs was abandoned, as it was deemed not feasible in several contexts. Instead, quasi-experimental methodologies, including difference-in-differences, instrumental variables, propensity score matching and pipeline approaches were applied according to the local

[9] The institutional dropout in different stages of the project had different reasons linked to external factors such as political instability and climate disasters, internal factors associated with human resources capacity and methodological issues related to the experimental methodologies (ILO, 2015).

conditions (ILO, 2015). The standardisation of the impact methodologies was replaced by more tailored approaches, even if difference-in-differences was applied in several of the individual projects.

In this brief presentation of reference projects in the context of the paper, a final mention goes to the evaluation project developed by Hulme and Mosley (1996) for the Overseas Development Administration (ODA) in the UK. Although this project adopted a summative evaluation approach, hence, a more conventional academic perspective on impact, the fact that the evaluations were implemented in different MFIs from five countries makes it an interesting case when analysing the potential standardisation of the evaluation process (or some of its elements).

The co-existence of different goals and approaches to impact assessment in microfinance is recognised by Hulme (2000), who advocates the development of impact monitoring capacity within the MFIs as a valid alternative to experimental scientific approaches, depending on the conditions of each project. In this perspective, he suggests that the effectiveness of an impact assessment should not be "automatically equated with the level of scientific proof that a study can claim", being important to evaluate the fit between its objectives, the financial and human resources involved and its context (Hulme, 2000, p.93).

3. The research design and Lendwithcare project

3.1. Research questions and methodologies

The research conducted follows from my personal participation in the impact assessment conducted by LWC of two microcredit programmes developed by its local partners. It is based on a case study methodology implemented in two settings (Pakistan and Zimbabwe).

The research question explored in the paper addresses the 'common project approach' adopted by LWC in the two settings, identifying its main elements and how these have influenced the results of the individual evaluations and the perception of the partners regarding their participation in the evaluations. This broad question was segmented in three sub-questions:

- *To which extent is it feasible to implement a common project approach to evaluation considering the diversity and complexity of the contexts in which microfinance programmes are implemented?*
- *What are the advantages and limitations of adopting a common project approach in the implementation of evaluations in different contexts?*
- *How can the application of a common project approach to evaluation contribute to the common knowledge and learning process of the microfinance and financial inclusion sector regarding its impact?*

The analysis is based on the cross-case synthesis of the two evaluations implemented, using the evidence collected through participation-observation and interviews with the evaluation team members.

Participant-observation creates ""unusual opportunities for collecting case study data", providing access to people and events and the opportunity to perceive reality from the perspective of an insider (Yin, 2018, p.124), as well as allowing for the development of detailed descriptions and in-depth analysis of the cases (Jorgensen, 1989). In this methodology, the researcher plays two different roles – one as participant/group member and other as researcher, studying the group from inside (Atkinson and Hammersley, 1994). While the direct involvement of the researcher/observer enhances the "opportunities to test understandings and meaning attribution" (Keiding, 2011, p. 108), it implies significant challenges. Among these are the time balance between the two roles and the possibility of bias in both recording and interpreting data (Yin, 2018). Reflexivity is, thus, important both in the application of the methodology (Atkinson & Hammersley, 1994; Gravensteijn, 2014; DeWalt & DeWalt, 2011) and in the identification of the limitations of the research.

The interviews with the evaluation team members were conducted after the implementation of the second wave of the household surveys in the two countries, with the objective of documenting their perspectives of the different evaluation team members regarding the evaluation process. Four management/staff members of the two MFIs (coded in the research as PAK1, PAK2, ZIM1, ZIM2) and two LWC team members (LWC1, LWC2) were interviewed between March and June 2018.[10] In the following sections of the paper, I refer to the evidence collected in the interviews through citations from the six interviewees (identified as personal communications or p.c.).

Before presenting the LWC project, I would like to stress that there was no intention of comparing the evaluation results of the two MFIs participating in the project, as they are not directly comparable. The design and implementation of any evaluation is highly influenced by contextual factors (House, 2003; Patton, 2008; Pawson and Tilley, 1997), which can affect all stages of the process. Hence, the idea of "one-size-fits-all" evaluation model is not applicable to the microfinance sector, and it does not correspond to the idea of common approach developed within the paper. The focus of the cross-case analysis was on the evaluation process itself, including the methodologies applied, and on the partnership established between LWC, the University of Portsmouth and the MFIs to deliver the project.

Having this in mind, it is important to define the meaning of 'common project approach' to evaluation in the context of the LWC project and the paper. LWC evaluation team, as leaders of the project, were aware of the advantages of a "tailored approach" (LWC1, personal communication, June 4, 2018), dependent on the country context and the programme evaluated. However, they were also keen in achieving a balance between quality and cost, rationalising as much as possible the scarce resources available for the project.

The standardisation associated with the 'common project approach' translated into:

- Similar overall research questions (and outcomes evaluated);

[10] The transcripts of the interviews were validated by each respondent, and data was analysed using Nvivo.

- Similar methodologies (longitudinal surveys; comparable sampling and data collection techniques);
- Evaluation teams with common elements (LWC, UoP);
- Same team structure at local level (involving a MFI manager and an operational project manager).

The data analysis, despite adjustments to accommodate the specificities of each dataset, included also common aspects, such as the exploration of heterogeneity among the microcredit clients. This approach did not prevent the need to adapt some aspects of the two evaluations to the particularities of the local contexts and resources of the two institutions, which were very different in terms of scale and capability to access resources. The adjustments included decisions regarding the sample size, the language of the questionnaires, the recruitment process of the interviewers, the local of the interviews, and the strategies to locate and contact the non-clients.

Considering that standardisation in any form can raise doubts among the readers, one of my goals in the paper is to describe the approach followed in the LWC project. For this reason, I give particular attention to identifying the feasibility conditions of the project (first research sub-question), and why the approach is worthy the attention of different actors involved in the sector.

3.2. Lendwithcare evaluation project

Lendwithcare is a crowdfunding platform created in 2010, by initiative of CARE International UK, a British development NGO. CARE has a long history in the provision of social programmes in developing countries, including financial inclusion through their village and savings loans programme (VSLA). LWC started as an alternative way to work financial inclusion through new technologies, particularly the novelty crowdfunding techniques. By the end of 2018, the platform worked in 11 countries; it had supported more than 95,000 low-income entrepreneurs (over 80% of whom are women), and lent $20 million through the provision of microcredit loans. There were around 50,000 active lenders profiles in the platform, mostly based in the UK.

The crowdfunding model followed by LWC relies in partnerships with local MFIs, which screen the applicants to the microfinance loans and undertake all the administrative process associated with the credit contract and repayment collection. The MFIs upload microfinance candidate profiles, who have their loans previously approved, to the platform. The money gathered from the crowd-lenders represents for the MFIs one of their funding sources. In the case of LWC, this funding is obtained at zero cost to the MFIs as no interest rate is charged and the lenders are the ones taking the credit risk. In addition, LWC helps the local partners with the operational costs associated with the participation in the platform.

Table 1 lists the local MFIs collaborating with LWC by the end of 2018. As it may be seen in the table, the institutions are diversified in experience (maturity), legal status, scale of the microcredit operation and repayment performance. All of them privilege social goals in their missions.

Table 1 – Lendwithcare Local Microfinance Partners

Country	Partner MFI	Established	Legal Status	No. Active Borrowers	Average loan size	PAR 30 days
Cambodia	BORVOR	2003 (as NGO)	NBFI (since 2017)	3,433	$2,828	2.40%
Ecuador	FACES	1991	NGO	17,340	$2,564	2.20%
Ecuador	Cooperativa Santa Anita	2001	Cooperative	3,596	$3,612	1.10%
Malawi	Microloan Foundation	2002	NGO	26,098	$81	4.60%
Pakistan	**Akhuwat (AIM)**	**2001 (as NGO)**	**NBFI (2017)**	**820,071**	**$144**	**0.26%**
Peru	IDERCV	1990	NGO	2,233	$860	3.92%
Philippines	LAMAC	1992	Cooperative	42,015	$400	9.20%
Rwanda	Umutanguha	2003	Cooperative	9,519	$520	5.60%
West Bank and Gaza	Reef Microfinance	2007	NGO	3,614	$3,512	6.60%
Vietnam	MACDI	2007	NGO	4,608	$221	1.7%
Zambia	Microloan Foundation	2008	NGO	8,409	$121	6.7%
Zimbabwe	**Thrive Microfinance**	**2012**	**NBFI**	**2,199**	**$655**	**9%**
Zimbabwe	Microloan Foundation	2016	NBFI	185	$124	0%

Source: Own construction from data available in the LWC website[11]
Note: Data reports to June 2018, except for IDERCV and Akhuwat Islamic Microfinance (2017)

The main goal of LWC is to "improve the lives of poor people" in developing countries (LWC1, p.c., June 4, 2018). However, during the first five years of the project, only anecdotal evidence of the outcomes of the microcredit programmes was gathered by "collecting case studies on an ad hoc way" (LWC2, p.c., June 4, 2018). Aiming to change this situation, LWC looked for an academic partner, which could support the design and implementation of an impact assessment project. The University of Portsmouth became this partner in the beginning of 2015.[12]

Equally, most of LWC local partners were not rigorously evaluating their social performance - in some cases, they had "really done anything" in this area (LWC1, p.c., June 4, 2018), or were doing

[11] Data retrieved from https://lendwithcare.org/info/how-it-works/microfinance-institutions, except data from Akhuwat (AIM), which was obtained during the implementation of the evaluation project.

[12] Despite the active involvement of the University of Portsmouth in the design and implementation of the evaluations, the leader of the project was LWC who ultimately took decisions taking into account the University recommendations and, when justified, the local MFIs opinions.

it incipiently.[13] This was the case of the two institutions selected by LWC to the project, Akhuwat Islamic Microfinance (AIM) and Thrive Microfinance. The selection of the institutions had also into account the high motivation and commitment of the MFIs' managers, the human resources capacity available for the project and the location in two English-speaking countries, which facilitated the communications between the teams.

The two MFIs, although different in many aspects, fit into the classification of Tier 2 and 3 proposed within the framework of the European Microfinance Platform. This classification includes criteria on size (measured in value of assets), sustainability (Return on Assets) and transparency (audited accounting).[14] In the beginning of 2015, Thrive was still a young MFI, not fulfilling any of the three criteria to be considered Tier 2 (Thrive, 2018). In turn, Akhuwat was a mature NGO, complying with the Tier 2 size and transparency criteria, but difficult to classify in terms of sustainability given the microcredit model implemented. Khan et al. (2017) estimate that the MFI's 'cost coverage ratio' in the period 2014-15 was 109%, but costs were covered mainly by donations as the institution does not charge interest rate in its loan products. In this way, it cannot be classified as sustainable using a conventional definition. Thus, for different reasons, both institutions corresponded to Tier 3 MFIs.

The impact assessment project implemented by LWC between 2015 and 2017 was based mainly on quantitative methodologies, namely a longitudinal survey to a sample of clients and non-clients of the institutions. The baseline surveys were implemented in 2015 in Pakistan and 2016 in Zimbabwe, while the second waves of the surveys were conducted in 2017, after the clients of both institutions had completed their first loan cycle.[15]

Sampling criteria were similar in both countries: in the branches directly working with LWC, i.e. uploading client profiles into the platform, all new clients with loans approved during the period of the survey were invited to participate in the project. Non-clients were randomly selected among entrepreneurs living/working in the same neighbourhoods, with similar characteristics to the clients but who did not have active business loans. Sample sizes were determined considering the number of active borrowers of the institutions as well as cost and logistic constraints.

Table 2 shows the sample distribution in the two surveys. In Pakistan, the clients were selected from four branches in Lahore and Kasur; these were among the network of 300 branches of the institution at the time, located mainly in the Punjab province. In Zimbabwe, the institution had

[13] Thrive Microfinance managers, from the MFI creation in 2012, adopted a pro-active attitude towards social performance assessment, and were taking the first steps in this direction when they were presented with the opportunity to participate in the LWC project.

[14] Tier 2 MFIs have between $5 and 50 million in assets, ROA positive in one of the last 3 years, and no less than -5% in the other years and audited financial statements for at least 3 years. Institutions not fulfilling all these criteria are classified as Tier 3 institutions (e-MFP, 2013).

[15] The first stage of the project between 2015 and 2017 included the implementation of two waves of the household surveys. LWC has decided to continue the project in 2019. In Pakistan and Zimbabwe, a third wave of the household surveys will be implemented; and the project will be extended to a third country, Ecuador, with the implementation of the baseline survey planned for the second semester of the year.

only two branches (Willowvale and Chitungwiza), both contributing to the crowdfunding platform and participating in the evaluation.

Table 2 – Household Survey Sample size

		Baseline	Survey 2017
AIM (Pakistan)	Clients	500	447
	Non-Clients	100	52
Thrive (Zimbabwe)	Clients	342	245
	Non-Clients	157	110

Source: Own construction

Following the data collection process, in which both MFIs were involved in some degree, the University of Portsmouth has independently conducted the data analysis and produced the evaluation reports.[16]

3.3. Case studies in Pakistan and Zimbabwe: similarities and differences

In order to better understand the standardisation elements and answer the research questions, it is helpful to identify the main similarities and differences between the two case studies. At first sight, the two evaluation settings could hardly be more different – two different continents, cultures, ethnicities and religions. Islam plays a major role in AIM microcredit programme, while in Zimbabwe where the majority of the population is Christian, religion is not a factor from an institutional perspective. From a macro perspective, Pakistan classifies as a lower-middle income country while Zimbabwe is among the low-income countries. Particularly important in the LWC project, the economic situation in Zimbabwe altered significantly during the study period, with increasing difficulties and uncertainty characterising the period. In contrast, Zimbabwe outperforms Pakistan in terms of literacy levels, which reflected in the evaluations - 91% of Thrive clients at baseline had attained secondary education or above, whereas in Pakistan 48% of the clients were illiterate and a further 16% had completed primary education.

Geography was also quite different, with strong implications in the design and field implementation of the surveys. Despite the branches selected in both settings being located in urban areas, population density and geographical dispersion were very different. In Pakistan, each branch had a relatively small radius of operations in areas highly populated, which facilitated the implementation of the surveys in terms of locating the respondents. In Zimbabwe, Harare and Chitungwiza are geographically dispersed, with comparatively lower population density. As a direct consequence, the interviews with clients were conducted at the homes/businesses of AIM clients, in the case of Pakistan, but at THRIVE offices in Zimbabwe, with clients being offered a small monetary compensation for their time and the cost of transport.

[16] Further information on the project and the evaluations conducted is available on the reports made available in the internet sites of the institutions.

There are also differences between the two case studies at the level of the institutional characteristics and the microcredit programmes implemented. At baseline, AIM was already a mature MFI, with a strong local focus, experienced staff and a highly formal hierarchical structure. During the research period, the MFI would grow significantly, becoming the largest microfinance institution in Pakistan in number of active borrowers (Khan et al., 2017). THRIVE is a small and young institution, founded by external investors, employing a capable but less experienced team, and with a comparatively more horizontal organisational structure.

These institutional differences translate into distinct perceptions of the MFIs' managers regarding evaluation, social performance management, and the participation of their institutions in the LWC project. Ultimately, the institutional culture contributed to the decision of AIM managers to develop their own internal evaluation mechanisms, along a new management information system. The participation in the LWC project has given strength to the idea that the MFI needs evidence on the social impact of its programme "not only for the researchers and external parties, but for the staff and for the board of directors" (AIM1, p.c., March 21, 2018). Internal legitimacy became noticeable among the MFI management priorities.

Equally, the institutional characteristics described help explain the focus of THRIVE managers on external legitimisation, their interest in the continuation of the evaluation project led and financially supported by LWC, and their enthusiasm towards externally validated evaluation tools, such as the Probability Poverty Index (PPI).

Finally, the two institutions implement microcredit programmes with different characteristics in terms of the credit product. AIM programme is an example of Islamic microfinance based on a credit-only product, no charge of interest (*qard hasan* loans), and a family approach, which implies the involvement of other family members in the loan process. Microcredit applicants might be male or female; by June 2016, 42% of the clients were women (Khan et al., 2017). Thrive illustrates the conventional interest-based microcredit model, adopting a microfinance plus approach, which includes compulsory savings and training prior to the loan approval. In the case of Thrive, all clients are women. The two MFIs differ as well in the average duration of the loan cycles, which is longer in AIM (18 months) relatively to Thrive (6 months).

These distinct characteristics of the programmes were in the LWC project more significant *per se* in the analysis of the individual evaluation results. In the design and implementation of the evaluation project, given the focus on main outcome indicators, these differences implied some adjustments, but they were not a deterrent to the 'common project approach'. In fact, the contextual, institutional and programme differences identified illustrate the features of the two settings that made them interesting cases to study the common project approach. The complexity associated with the contextual and institutional factors in the two countries have introduced challenges in the implementation of the project, which allowed for testing the application of the common project approach in the emergence of unforeseen events.

Albeit these differences, the two case studies present as well similarities, which were fundamental to the positive views of all the partners regarding the approach followed in the LWC project. Among the similarities between the two cases, there are factors related to the status of financial inclusion and the microfinance sector in the countries. The microfinance markets in Pakistan and Zimbabwe, although in very different stages of development, are both growing, with increasing competition but still large segments of the population underserved, especially in the rural areas of the two countries.

According to the results of the latest Global FINDEX Survey, the overall access to formal finance in Pakistan and Zimbabwe has substantially improved in the period 2014-2017, however, this progress was accomplished through the growth of bank accounts (mobile accounts in the case of Zimbabwe), not reflecting in the same way in the credit data. In Pakistan, there was a slight improvement in the access to formal credit by the poorest segments of the population, which seems to reflect the increasing dynamism of the microcredit sector. However, access to finance remains at low levels and well below the average numbers for the countries with similar economic development levels (World Bank, 2018). Regarding external factors, the two countries share a negative feature related to political instability, which aggravates the effects of other external events such as the cash crisis in Zimbabwe, the terrorist attacks in Pakistan or the extreme climate episodes that have affected both countries in recent years.

From an institutional perspective, both microcredit programmes are based on group lending methodologies and aspire to have social impact, improving the lives of the clients by giving them access to business loans to develop their entrepreneurial activities. This shared social focus translates into an interest in assessing the performance of the institutions regarding a range of common outcomes, which facilitated the task of developing a similar evaluation methodology. Many of the questions included in the surveys were the same, or they were slightly modified in the hypotheses of answer, with the differences between the two cases being established in the interpretation of the results.

As mentioned above, both MFIs were in the beginning of the project at an incipient stage in terms of social performance measurement, even if Thrive had taken some steps in this area.[17] They had also management information systems not much sophisticated, with AIM branches working on a paper basis and Thrive database not providing individual loan data for the members of the solidarity groups. Adding to this, there was also some unfamiliarity with the evaluation process.

The similarities mentioned so far were significant, but probably the most important factor present in the two settings, which contributed to the successful implementation of the evaluations, was the firm commitment of the MFI managers throughout the project. This strong motivation was initially triggered by legitimisation purposes – the need to produce credible evidence of the outcomes of the programmes, which guaranteed the buy-in by the management of the two institutions. This motivation and commitment was also crucial at the level of the local evaluation teams, which had no alterations during the project. Common to all those involved in the LWC project with decision-making power over the programmes was a perspective of utility

[17] Including the creation of a dedicated staff position (social performance manager), the introduction of client satisfaction surveys and the initial steps in the development of an internal poverty assessment tool (since there was no externally validated poverty index, such as the PPI developed for Zimbabwe).

of the evaluations. This was conveyed by LWC1 in the final interview when considering that the LWC team would have not implemented the project if it was not clear “from the beginning that this is of practical use to all parties involved” (LWC1, p.c., June 4, 2018).

4. Discussion of the findings

4.1. Standardisation in the LWC assessment project: the common project approach

As presented above, there are three main areas where standardisation played some role in the LWC project:

- Evaluation overall goals and outcomes of interest;
- Methodological choices (longitudinal household surveys; similar sampling criteria; data collection and analysis tools);
- Human resources (same core evaluation team, similar structure of the local teams and recruitment criteria for the survey enumerators).

In comparison with the LWC project, the approach followed in the AIMS project appears to be a good benchmark as the three core evaluations were implemented in urban areas located in different continents, and the methodologies applied seemed to be generally similar, including longitudinal data, control group formed by non-clients, and data analysis using a before-after comparison (Goldberg, 2005). The major difference in the design of the evaluations between the two projects relates to the client sample. In the LWC project, the clients participating in the first round of the survey were new clients, which allowed for collecting baseline data, while in the AIMS project, they were existing clients in different loan cycles.

A common methodological approach was also adopted in the evaluations conducted by Hulme and Mosley (1996). Sebstad and Chen (1996, p.4) suggested that this “consistency in the methods”, with individual studies sharing the “same basic framework, set of research questions, and, in most cases, impact variables”, was useful in producing substantial evidence in relation to the research questions and allowing for addressing key policy issues.

In contrast, the LWC project seems to differ from the Microfinance for Decent Work and the Imp-Act projects in several ways. The scale of these projects implied much more diversity in terms of the contexts and the institutions involved. Consequently, in these two projects there was a more customised methodological approach, with different methodologies being implemented in response to differentiated research questions and implementation challenges. Despite these differences, there were somepoints of contact between the individual projects, as in the case of the adoption of similar AIMS tools in many of the Imp-act projects, but there was not a common approach within the projects in this regard.

In this comparison, the projects scale and scope are crucial elements. Not only the number of participating MFIs is significantly higher (30 in the Imp-Act project and 13 in the ILO project), also the academic coordination team (in both cases located in the UK) and the local teams represented a level of costs that is not comparable with the investment of LWC in this first phase of its project. This should not be surprising if considering the size of LWC as an institution (or rather said as a small department within a large development organisation) and the adopted exploratory and learning approach to impact evaluation and social performance assessment.

4.2. Cross-case analysis: answering the research question(s)

The results and lessons from my experience as academic consultant in the LWC project are presented with reference to the three research (sub)-questions introduced in section 3.

The feasibility conditions

There were different elements contributing to the positive perception of LWC and the two local partners regarding the common project approach. Although it is difficult to assert the weight of each of these factors, I suggest that all of them played a part on the valorisation of social performance assessment within the institutions.

One of these factors is the active participation in the two settings of the core LWC team and me as the academic consultant, which guaranteed consistency in the implementation of the project and the permanent exchange of experiences between partners in all moments. This approach compares with the common lead assumed by Hulme and Mosley in the ODA project (Hulme & Mosley, 1996), but contrasts with the approach followed in the AIMS project. In this case, the team of evaluators seem to have been different in the different studies, despite interactions between the researchers within the global AIMS project (Barnes, 2001; Chen & Snodgrass, 2001; Dunn & Arbuckle Jr, 2001). In the Imp-Act and ILO projects, the approach seemed to have be similar to the AIMS project. Furthermore, in any of these four projects there is a mention to the direct involvement of the funders of the project/evaluations in the implementation of the impact assessments. Copestake and Simanowitz (2005) refer to the participation of the Ford Foundation in the initial stages when discussing the project as a whole, but not in the individual projects at the MFIs level.

Other important element was the association with an academic partner, which was perceived as increasing the impartiality of the process and the credibility of the evaluation results. This was one of the reasons presented by the LWC manager to look for an academic consultant - “we have a lot of credibility of the research, the way it was done, who has done it” (LWC1, p.c., June 4, 2018). This idea was corroborated by the Thrive manager, considering that if ‘you’ collaborate with “a reputable development partner and an academic/research institution, your findings are obviously authoritative and can be accepted as credible” (THR1, p.c., May 18, 2018).

A third important factor related to the motivation and commitment of the participating MFIs, which were willing to implement the evaluations and to develop capacity within the institutions, but that at the start of the project had not gathered the conditions (motivation, leadership, skills, funds) to implement social performance assessment and impact evaluation in a more systematic and rigorous way. The starting point of the Tier 3 institutions was, thus, relevant for the results.

Finally, it was fundamental the experimental approach and learning attitude of the LWC team, the evaluator and the field partners in regard to the evaluation project, which reflected in the initial focus on broad categories of outcomes. The quantitative surveys focused on general questions regarding the outcomes of the programmes at the client level, requiring some adaptation to the local conditions but still allowing for the application of a common set of questions in the two cases.

Advantages and limitations of the approach

The approach presents both advantages and limitations, which are different considering the different partners. From the perspective of the LWC team, the approach allowed for accessing additional information on the two MFIs, harmonised to a certain extent as similar indicators were used for the broader categories of outcomes. It equally permitted to build technical capacity within the local partners on evaluation and social performance assessment. The implementation of the impact project, and its potential extension to other LWC field partners, can therefore contribute to build communalities across the LWC portfolio and strengthen the partners' internal capacity.

Furthermore, these advantages were achieved using common resources, including the academic consultancy, improving the cost-effectiveness associated with the evaluations. LWC1 considered in the final interview that "we have done this at a very low cost in terms of the research we do generally in CARE" (LWC1, June 4, 2018). There were certainly cost benefits in this approach, but it should be noted that part of the cost reduction in this phase of the project was due to the direct contribution of the University of Portsmouth, by supporting my work through a PhD bursary.

To the MFIs participating in the project, there were two main advantages. First, they had the opportunity to participate in a multiple-country impact evaluation project led by LWC, an initiative of CARE International UK, a "reputable development partner" (THR1, p.c., May 18, 2018) working in 74 countries. In terms of the credibility of the results and the dissemination of the work conducted locally by the MFI, the project opened possibilities that were hardly achievable within an independent evaluation project. Second, the application of common methodologies in the two settings allowed for an enriched learning process, integrating experiences and knowledge from different contexts and different perspectives, while pursuing common overall evaluation goals.

Equally significant in the LWC project was the promotion of a shared understanding of evaluation, and the development of a common language among the partners, both contributing to gradually change the mind-set of the MFI managers towards evaluation. This was mentioned by the LWC manager: "our partners are now beginning to take the issue of evaluation research much more seriously as well" (LWC1, p.c., June 4, 2018). Finally, and as a cumulative result of these identified advantages, the common project approach had associated a motivational factor, enhanced by the partners shared goal of improving the lives of poor populations in their respective countries. The advantages mentioned resulted, in part, from the shared core evaluation team, which was possible given the small scale of the LWC project. It would have been hardly achievable in large projects such as Imp-Act or Microfinance for Decent Work. Still, considering the responsibilities of the members involved, there were constraints in terms of the timings and availability of the project common members, located in the UK, to work in the project and to travel to the two countries. The fact that I was a full-time PhD during the implementation of this phase of the project was helpful, as it gave me some flexibility to adapt to the timings of the other partners.

More important than these operational issues, the characteristics of the partnership established in what concerns the relations between the partners, particularly between LWC (leader and funder of the evaluations) and the local institutions, should be considered carefully. The two case studies conducted illustrate two very different situations in this regard. The MFI in Pakistan was a mature institution, growing significantly and not dependent of the funding obtained from LWC. The partner in Zimbabwe was a small, young MFI, trying to achieve financial sustainability in a challenging environment, hence, in need to present credible evidence of its results to maintain the existing funders (including LWC) and attract new funders. Considering the possibility of extension of the project to other LWC field partners, other scenarios may emerge, being important to pay attention also to the interactions between the MFIs, which are very different in size, maturity, programmes implemented, etc.

One final limitation of the approach followed in the LWC project respects to the broad perspective taken on the outcomes of the programmes to be evaluated. A more in-depth analysis of the programmes, evaluating the programmes' processes and mechanisms of change associated with the interventions and taking a dynamic approach to the social impact of the programmes at different levels would imply the adoption of different evaluation models, including Guba and Lincoln's (1989) constructivist model or Pawson and Tilley's (1997) realistic. It is, although, important to have in mind that these approaches are associated with the implementation of different methodologies or combinations of methodologies (mixed methods) and require a different set of resources (financial, human, time), which were not available in the initial stages of the LWC project

Contributions to the microfinance sector

Simanowitz (2003) attributed an important role to funders in the support of research on impact and social performance, especially for MFIs with limited resources. In line with the approach followed in the Imp-Act and Microfinance for Decent Work projects, the LWC project did not aim to produce summative judgment evaluations *per se* (Patton, 2008), actively seeking the involvement of LWC own team and the MFIs in the different stages of the evaluation process.[18] This approach loses in the independence of the findings, and thus, on its utility as a pure judgmental instrument, but gains (enormously) in the development of evaluative thinking within the institutions and in building internal capacity in regard to evaluation and social performance assessment and management.

By being closer to the interests and objectives of the microfinance institutions, this approach can contribute to overcome the initial barriers faced by many MFIs, such as THRIVE, which have the interest but do not have the know-how or the required expertise, and do not have the necessary

[18] This level of involvement of the local partners differed according to the stage of the evaluation process and the different field challenges encountered in the two countries. In both cases, there was no participation of the LWC or the MFIs teams in the analysis of the data collected through the surveys and the production of the project reports, which were prepared by me as evaluator/academic consultant.

funds. At the same time, it can help minimise potential suspicious towards evaluation and external evaluation experts.

The contribution of the LWC project and its 'common project approach' in the two settings should be understood as one additional effort, among other initiatives, aiming to minimise objections and build a common understanding and language regarding evaluation, social performance management and impact assessment in the sector, which will ultimately contribute to value the work of different actors.

The AIMS project was one of these initiatives. Barnes and Sebstad (2000) suggested standards for credible, useful and cost-effective impact assessments in microfinance, taking into consideration that the academic standards for a rigorous impact assessment, which the authors associated with the implementation of costly experimental methodologies, were out of reach for the majority of the MFIs. As previously referred to the AIMS guidelines and tools were the methodological basis used by most MFIs participating in the Imp-Act action-research project. This choice was presented by Greeley (2005, pp. 48-49) as relating not only to the level of rigour associated with the tools, but also because they were considered to be "practical, cost-effective and relevant". This was important since in many of the cases, the MFIs had privileged the learning agenda regarding social performance in comparison with external accountability objectives.

This distinction of standards between a more academic approach (emphasizing rigour), and a more practitioner perspective (stressing usefulness), was acknowledged by the LWC manager who viewed academics as privileging maximum rigour while MFI managers and social investors were more flexible. This greater flexibility translated into willingness to accept lower standards resulting from compromises at field level in the implementation of the evaluations as long as these were deemed necessary to guarantee the feasibility of the assessment, and an adequate level of quality was ensured (LWC1, p.c., June 4, 2018).

Adopting a more plural perspective, this argument seems reasonable, but it raises the difficult question of identifying what should be considered as necessary compromises in terms of implementation and adequate level of quality. As the ILO report shows, any attempt to classify different evaluations in terms of quality is complex and can influence the analysis of the merits and limitations of the evaluations in a way that valid and useful information on the institutions and the programmes may be overlooked due to the perceived lower value of the process. Again, this does not mean that the different levels of quality should not be recognised, but that there will be probably lessons to be taken from most projects.

At this level, the work of supranational organisations, in this case SPTF, can make a difference by integrating the knowledge and practical insights from the different initiatives and projects in the development of harmonised standards and language for evaluation in the microfinance sector.

5. Conclusions

In a context of diversification of the actors and scope of microfinance, the focus on assessing social performance and social impact continues to be a priority in the sector. A lot has been done in the past two decades, both at academic level with the development of more sophisticated impact evaluation methodologies and at practitioner level with the dissemination of practical mechanisms and tools that allow field level institutions to measure and manage their social performance.

Still, there is a lot to be done, particularly in what concerns Tier 2 and 3 MFIs and new actors in the field (e.g. mobile operators). These institutions may have not been engaged with the networks and funders/donors which have been encouraging social performance management (such as SPTF), or they might not have the financial and human resources to kick-start the process or they might just be discouraged by lack of familiarity with the process. Truly, in many of these cases, formal impact assessments as defined by Copestake (2012) are not a feasible option, but this does not mean that the institutions cannot and should not measure their social performance and impact, considering here the broader concept of impact identified by White (2009). As referred to throughout the paper, fundamental in any situation is to be transparent regarding the objectives, definitions and approach adopted, allowing the audience of the project to make their own decision on the quality and utility of the information provided.

This idea is not new as illustrated by the references to the AIMS, Imp-Act and Microfinance for Decent Work projects (Barnes and Sebstad, 2000; Copestake and Simanowitz, 2005; ILO, 2015). In the beginning of 2000s (already before the emergence of many of the sector delinquency crises), Hulme has called attention to an alternative to scientific (academic) impact evaluation focusing on the internal social impact and performance assessment mechanisms within the MFIs. This idea was endorsed as well by Simanowitz (2001), who in a later paper refers to the prominent role to be played by donors and funders in this process (Simanowitz, 2003).

The LWC project fits in these proposals. The reduced scale of the project, the characteristics of the participant MFIs (initially classified as Tier 3), and the direct involvement of the LWC team in the field implementation of the evaluations, give to the project some originality relatively to the three other projects, adding to the cumulated knowledge on the topic. Even if from a purely academic perspective, the methodological limitations associated with the quasi-experimental design and field challenges preclude the corroboration of attribution of the observed effects to the participation in the microcredit programme, the evaluations in the two settings provided credible evidence on a number of outcomes for different stakeholders, internal and external. The LWC project showcases an evaluation approach well received by the MFIs involved and raising interest on other LWC local partners. Albeit some of its features, particularly the active participation of the programme stakeholders, limit the level of independency of the process, this type of approach allows for gaining commitment towards evaluation and the use of its findings and for developing evaluative thinking within the institutions, making the case for social performance management in MFIs privileging social goals. The common project approach adopted by LWC reinforced these gains by promoting the exchange of experiences between all the participants.

Moreover, although the LWC project involves traditional microfinance providers,[19] it would be reasonable to think that many of the new actors entering the microfinance market, providing financial services to poor and vulnerable populations, such as mobile operators, will follow principles of no harm to their clients. Some may be more ambitious regarding the social impact of their products, even if not motivated by explicit social missions. The dissemination of experiences such as the LWC project can, hence, be helpful for these new actors and contribute to improve social performance in the whole sector.

References

Adams, D.W. & Vogel, R.C. (2013). Through the Thicket of Credit Impact Assessments. In J-P. Gueyie, R. Mannos & J. Yaron (Eds.), *Microfinance in Developing Countries – Issues, Policies and Performance Evaluation* (36-61). Basingstoke: Palgrave MacMillan.

Afonso, J.S., Morvant-Roux, S., Guérin, I. & Forcela, D. (2017). Doing Good by Doing Well? Microfinance, Self-Regulation and Borrower's Over-Indebtedness in the Dominican Republic*. Journal of International Development*, 29 (7), 919 – 935.

Atkinson, P. & Hammersley, M. (1994). Ethnography and participant observation. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 248-261). Thousand Oaks, CA: Sage Publications.

Balkenhol, B. (2012). Impact Assessments in Microfinance: Theory and Practice. In *UMM Workshop Report - Microfinance in Crisis? Impact and Financial Transparency* (pp. 20-28). Luxembourg: European *Microfinance* Platform.

Banerjee, A. & Duflo, E. (2011). *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*. New York: Public Affairs.

Banarjee, A., Karlan, D. & Zinman, J. (2015). Six Randomized Evaluations of Microcredit: Introduction and Further Steps. *American Economic Journal: Applied Economics,* 7(1), 1-21.Barnes, C. (2001). Microfinance Program Clients and Impact: An Assessment of Zambuko Trust, Zimbabwe. Washington DC: Management Systems International.

Barnes, C. (2001). *Microfinance Program Clients and Impact: An Assessment of Zambuko Trust, Zimbabwe*. Washington DC: Management Systems International.

Barnes, C. & Sebstad, J. (2000). Guidelines for Microfinance Impact Assessments (Discussion Paper). Washington DC: Management Systems International

Bastiaensen, J., Marchetti, P., Mendoza, R., & Pérez, F. (2013). After the Nicaraguan Non-Payment Crisis: Alternatives to Microfinance Narcissism. *Development and Change*, 44(4), 861–885.

Bateman, M. (2010). *Why Doesn't Microfinance Work? The destructive rise of local neoliberalism.* London: Zed Books.

Bédécarrats, F., Guérin, I. & Roubaud, F. (2015). *The gold standard for randomized evaluation: from discussion of method to political economics* (CEB Working Paper 15/009). Brussels: Centre Emile Bernheim.

Boruch, R. F. (1997). *Randomized Experiments for Planning and Evaluation*. Thousand Oaks: SAGE Publications.

[19] Both institutions are presently classified as non-banking financial institutions (NBFI), after Akhuwat transformed from NGO into a regulated microfinance institution in 2017. The transformation covered the microcredit operation of the former NGO, becoming Akhuwat Islamic Microfinance (AIM). The non-financial services provided remained within the NGO (Khan et al., 2017).

Burki, H-B. (2009). *Unravelling the Delinquency Problem (2008/2009) in Punjab – Pakistan* (MicroNote No 10). Islamabad: Pakistan Microfinance Network.
Bylander, M., Res, P., Jacoby, L., Bradley, P., & Pérez, A.B. (2018). Over-indebtedness and microcredit in Cambodia: Moving beyond borrower-centric frames. *Development Policy Review*, doi:10.1111/dpr.12399
Chen, G., Rasmussen, S. & Reille, X. (2010). *Growth and Vulnerabilities in Microfinance* (CGAP Focus Note No 61). Washington DC: CGAP.
Chen, M.A. & Snodgrass, D. (2001). Managing Resources, Activities, and Risk in Urban India: The Impact of SEWA Bank. Washington DC: Management Systems International
Copestake, J. (2012). The Heterogeneity of Microfinance Impact and its Assessment. In *UMM Workshop Report - Microfinance in Crisis? Impact and Financial Transparency* (pp.8-14). Luxembourg: European Microfinance Platform.
Copestake, J. & Simanowitz, A. (2005). Introduction. In J. Copestake, M. Greeley, S. Johnson, N. Kabeer, A. Simanowitz & K. Knotts (Eds.), *Money with a Mission (Volume 1): Microfinance and Poverty Reduction*. UK: ITDG Publishing.
DeWalt, K. M. & DeWalt, B. R. (2011). *Participant observation: a guide for fieldworkers* (2nd Edition). Lanham, Maryland: AltaMira Press.
Duflo, E. (2004). Scaling Up and Evaluation. In F. Bourguignon & B. Pleskovic (Eds.), *Annual World Bank Conference on Development Economics 2004: Accelerating Development (Bangalore Conference Proceedings)* (pp.341-369). Washington DC: The World Bank
Dunn, E. (1999). *Microfinance Clients in Lima, Peru: Baseline Report for AIMS Core Impact Assessment.* Washington DC: Management Systems International.
Dunn, E. & Arbuckle Jr, J.G. (2001). Microcredit and Microenterprise Performance: Impact Evidence from Peru. *Small Enterprise Development*, 12 (4), 22 – 33.
Duvendack, M., Palmer-Jones, R., Copestake, J., Hooper, L., Loke, Y. & Rao, N. (2011). *What is the evidence of the impact of microfinance on the well-being of poor people?* London: EPPI Centre, Social Science Research Unit, Institute of Education, University of London.
e-MFP (2013). *Working towards a common consensus on the definition of Tiers in microfinance* (Discussion Paper 1, April 2013). European Microfinance Platform: Luxembourg.
Fernando, J. (1997). *Non-governmental Organizations, Micro-Credit, and Empowerment of Women.* The Annals of the American Academy of Political and Social Science, 554 (1), 150 – 177.
Gertler, P., Martinez, S., Premand, P., Rawlings, L. B. & Vermeersch, C. (2016). *Impact Evaluation in Practice (*2nd Edition). Washington DC: The World Bank.
Goldberg, N. (2005). *Measuring the impact of microfinance: Taking stock of what we know.* Washington, D.C.: Grameen Foundation.
Gonzalez, A. (2008) *Microfinance, Incentives to Repay, and Overindebtedness: Evidence from a Household Survey in Bolivia* (PhD Dissertation). The Ohio State University, Columbus.
Gravesteijn, R. (2014). *Models of Social Enterprise? Microfinance Organisations as Promoters of Decent Work in Central Asia* (PhD Dissertation, Volume 1). University of Bath, UK.
Greeley, M. (2005). Direct Material Impacts. In J. Copestake, M. Greeley, S. Johnson, N. Kabeer, A. Simanowitz & K. Knotts (Eds.), *Money with a Mission (Volume 1): Microfinance and Poverty Reduction* (pp. 46-65). UK: ITDG Publishing
Guba, E.G. & Lincoln, Y.S. (1989). *Fourth Generation Evaluation*. Newbury Park: SAGE Publications.
Guérin, I., Kumar, S. & Agier, I. (2013). Women's Empowerment: Power to Act or Power Over other Women? Lessons from Indian Microfinance. *Oxford Development Studies*, 41 (sup.1), 76 – 94.

Guérin, I., Labie, M & Servet, J.M. (2015). Introduction. In I. Guérin, M. Labie & J. M. Servet (Eds)., *The Crises of Microcredit* (pp. 1-20). London: Zed Books.
House, E. R. (2003). Stakeholder Bias. *New Directions for Evaluation*, 97, 53-56.
Hughes, K. & Hutchings, C. (2011). *Can we obtain the required rigour without randomization? Oxfam GB's non-experimental Global Performance Framework (*3ie Working Paper 13). New Delhi: International Initiative for Impact Evaluation.
Hulme, D. (2000). Impact Assessment Methodologies for Microfinance. *World Development*, 28 (1), 79-98.
Hulme, D. & Mosley, P. (1996). *Finance Against Poverty: Volume 1*. London: Routledge.
International Labour Organisation (2015). *Microfinance for Decent Work*. Geneva: ILO.
Jorgensen, D.L. (1989). The Methodology of Participant Observation. In D.L. Jorgensen (Ed.). *Applied Social Research Methods: Participant Observation* (pp.12–25). Thousand Oaks, CA: SAGE Publications.
Kabeer, N. (2001). Conflicts Over Credit: Re-Evaluating the Empowerment Potential of Loans to Women in Rural Bangladesh. *World Development*, 29 (1), 63-84.
Karlan, D., Goldberg, N. & Copestake, J. (2009). Randomized control trials are the best way to measure impact of microfinance programs and improve microfinance products design. *Enterprise Development and Microfinance*, 20 (3), 167 - 176.
Keiding, T. B. (2011). Observing Participating Observation – a Re-description Based on Systems Theory. *Historical Social Research*, 36 (1), 106 – 123.
Khan, A.A., Ishaq, M.S., Afonso, J.S. & Akram, S. (2017). Is it possible to provide *qard hasan* and achieve financial self-sustainability? The experience of Akhuwat in Pakistan. In M. Harper & A.A. Khan (Eds), *Islamic Microfinance – Shari'ah compliant and sustainable?* Rugby, UK: Practical Action Publishing.
Littlefield, E., Morduch, J., & Hashemi, S. (2003). *Is microfinance an effective strategy to reach the millennium development goals?* (CGAP Focus Note 24). Washington DC: CGAP.
Mader, P. (2013). Rise and fall of microfinance in India: The Andhra Pradesh Crisis in Perspective. *Strategic Change*, 22, 47-66.
Marr, A. & Awaworyi, S. (2012). Microfinance Social Performance: A Global Empirical Study. *Applied Econometrics and International Development,* 12 (2), 51 – 68.
Morduch, J. (1998, June). *Does Microfinance Really Help the Poor? New Evidence from Flagship Programs in Bangladesh* (Unpublished Paper).
Morduch, J. (1999). The Microfinance Promise. *Journal of Economic Literature,* 37, 1569 – 1614.
Morvant-Roux, S., Afonso, J., Forcella, D. & Guérin I. (2015). How good repayment performances can harm borrowers: evidence from Dominican Republic. In I. Guérin, M. Labie and J.M. Servet (Eds.), *The Crises of Microcredit* (pp. 92-112). London: Zed Books.
Morvant-Roux, S. & Roesch, M. (2015). The social credibility of microcredit in Morocco after the default crisis. In I. Guérin, M. Labie & J. M. Servet (Eds)., *The Crises of Microcredit*. London: Zed Books.
Nair, T. (2011). Microfinance: Lessons from a Crisis. *Economic and Political Weekly*, 46 (6), 23-26.
Odell, K. (2015). *Measuring the Impact of Microfinance – Looking to the Future*. Washington DC: Grameen Foundation.
Patton, M.Q. (2008). *Utilisation-Focused Evaluation* (4th edition). Thousand Oaks: SAGE Publications.
Pawson, R. & Tilley, N. (1997). *Realistic Evaluation*. London: SAGE Publications.
Pitt, M. & Khandker, S. (1998). The Impact of Group-based Credit Programmes on Poor Households in Bangladesh: Does the Gender of Participants Matter? *Journal of Political Economy,* 106 (5), 958 – 996.

Ravallion, M. (2012). Fighting Poverty One Experiment at a Time: A Review of Abhijit Banerjee and Esther Duflo's Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty. *Journal of Economic Literature,* 50 (1), 103 – 114.
Reichardt, C. S. (2011). Evaluating Methods for Estimating Program Effects. *American Journal of Evaluation,* 32 (2), 246 – 272.
Rhyne, E. (2001). *Commercialization and Crisis in Bolivian Microfinance*. Bethesda, USA: Development Alternatives, Inc.
Rogers, P. (2014). *Theory of Change* (Methodological Brief Impact Evaluation No. 2). Florence: UNICEF.
Rozas, D. (2011). *Weathering the Storm: Hazards, Beacons and Life Rats Lessons in Microfinance Crisis Survival from Those Who Have Been There*. London: Centre for Financial Inclusion.
Schicks, J. (2013). The Definition and Causes of Microfinance Over-Indebtedness: a Customer Protection Point of View. *Oxford Development Studies,* 41 suppl, 95 – 116.
Schicks J., & Rosenberg, R. (2011). *Too Much Microcredit? A Survey of the Evidence on Over-Indebtedness* (CGAP Occasional Paper19). Washington DC: CGAP.
Scriven, M. (2008). A Summative Evaluation of RCT Methodology: & An Alternative Approach to Causal Research. *Journal of Multidisciplinary Evaluation,* 5 (9), 11 – 24.
Sebstad, J. & Chen, G. (1996). *Overview of Studies on the Impact of Microenterprise Credit* (Report). Washington DC: Management Systems International.
Simanowitz, SA. (2001). From event to process: current trends in microfinance impact assessment. *Small Enterprise Development*, 12 (4), 11 – 21.
Simanowitz, A. (2003). *Appraising the Poverty Outreach of Microfinance: A Review of the CGAP Poverty Assessment Tool* (Occasional Paper No 1). Brighton: Imp-Act
Smits, J. (2017). *Overindebtedness Among the Poor: Measurement, Causes and Policies* (PhD Dissertation). ETH Zürich
Van Rooyen, C., Stewart, R. & de Wet, T. (2012). The Impact of Microfinance in Sub-Saharan Africa: A Systematic Review of the Evidence. *World Development,* 40 (11), 2249 – 2262.
Vogelgesang, U. (2003). Microfinance in Times of Crisis: The Effects of Competition, Rising Indebtedness, and Economic Crisis on Repayment Behavior. *World Development*, 31 (12), 2085 – 2114.
Von Pischke, J.D. & Adams D.W. (1980). Fungibility and the Design and Evaluation of Agricultural Credit Projects. *American Journal of Agricultural Economics*, 62 (4), 719-726.
White, H. (2009). *Theory-Based Impact Evaluation: Principles and Practice* (3ie Working Paper 3). New Delhi: International Initiative for Impact Evaluation.
World Bank (2018). *The Little Data Book on Financial Inclusion 2018*. Washington DC: World Bank.
Wright, K. & Copestake, J. (2004). Impact Assessment of Microfinance Using Qualitative Data: Communicating between Social Scientists and Practitioners Using the QUIP. *Journal of International Development*, 16, 355-367.
Yin, R. K. (2018). *Case Study Research and Applications: Design and Methods* (6th Edition). Thousand Oaks: SAGE Publications.